

# Platform-Driven Collaboration Patterns: Structural Evolution Over Time and Scale

Negin Maddah  and Babak Heydari 

**Abstract**—Within an increasingly digitalized organizational landscape, this research explores the dynamics of decentralized collaboration, contrasting it with traditional collaboration models. An effective capturing of high-level collaborations (beyond direct messages) is introduced as the network construction methodology including both temporal and content dimensions of user collaborations—an alternating timed interaction (ATI) metric as the first aspect, and a quantitative strategy of thematic similarity as the second aspect. This study validates three hypotheses that collectively underscore the complexities of digital team dynamics within sociotechnical systems. First, it establishes the significant influence of problem context on team structures in work environments. Second, the study reveals specific evolving patterns of team structures on digital platforms concerning team size and problem maturity. Last, it identifies substantial differences in team structure patterns between digital platforms and traditional organizational settings, underscoring the unexplored nature of digital collaboration dynamics. Focusing on Wikipedia’s co-creation teams as a representative online platform, this study is instrumental for organizations navigating the digital era by identifying opportunities and challenges for managing information flow. The findings reveal significant collaborative potential and innovation in large online teams: the high speed of knowledge-sharing, numerous subcommunities, and highly decentralized leadership. This study paves the way for platform governors to design strategic interventions, tailored for different problem types, to optimize digital team dynamics and align them to broader organizational goals.

**Index Terms**—Decentralized collaboration, digital platforms, knowledge-sharing (KS), social networks, Wikipedia editorial network.

## I. INTRODUCTION

ONLINE platforms play a crucial role in today’s knowledge-sharing (KS) within sociotechnical systems. Peer production teams within these digital platforms illustrate a shift toward shared leadership [1] for knowledge-creating, a trend widely adopted in organizations for open innovation

projects [2], reducing the relevance of traditional hierarchies. Exploring the information flow within these decentralized structures is essential for digital platform governors. This study utilizes computational network analysis to explore how online teams behave in terms of KS flow, ideas diffusion speed, subcommunity formation, and decentralized leadership. By examining these network structure metrics across various stages of an artifact, the study seeks to provide insights into the complexities of online collaboration evolution [3]. We pose critical questions about the influence of problem context on team structure dynamics and the presence of any distinct structural patterns in the evolution of online teams.

Additionally, we explore the differences between the interaction networks of “Digital Platform-Driven Collaboration” and “Traditional/Hierarchical Organizational Collaboration.” The intricacies of digital decentralized collaboration are explored using Wikipedia co-creation teams [4], leveraging Wikipedia’s extensive data on diverse article topics, maturity levels, and collaboration team scales. On the other hand, the traditional collaboration reference is provided by the findings of previous studies.

Our research enhances the understanding of digital collaboration dynamics. It also introduces an innovative network construction methodology for capturing co-creation among collaborators in two dimensions of their work: temporal and content. Our approach provides insights into the complex behaviors and patterns that emerge in decentralized settings, enriching the academic discourse on digital peer production. Furthermore, by highlighting the differences between digital platforms and traditional organizational models, our findings are pivotal for organizations embracing digital platforms, guiding them toward more effective team structures and enhanced creativity. Utilizing the comparative findings of this study opens new directions for studying when centralized policymaking and interventions can be beneficial in sociotechnical systems design [5].

## II. LITERATURE REVIEW

In recent years, the landscape of KS within teams has undergone a significant transformation, driven by advancements in digital technologies [6], [7]. Digital platforms, functioning as sociotechnical systems that integrate social communities with technical infrastructure [8], [9], have transcended traditional geographical limits, becoming pivotal to innovation and design processes [10], [11], [12], [13], [14]. These platforms are not

Manuscript received 14 February 2024; revised 8 July 2024; accepted 27 August 2024. Date of publication 24 September 2024; date of current version 3 December 2024. (Corresponding author: Babak Heydari.)

Negin Maddah is with the Department of Mechanical and Industrial Engineering, Northeastern University, Boston, MA 02115 USA.

Babak Heydari is with the Department of Mechanical and Industrial Engineering, Network Science Institute, Northeastern University, Boston, MA 02115 USA (e-mail: b.heydari@northeastern.edu).

This article has supplementary downloadable material available at <https://doi.org/10.1109/TCSS.2024.3452028>, provided by the authors.

Digital Object Identifier 10.1109/TCSS.2024.3452028

merely tools but the core of shaping and governing the processes of knowledge creation. This shift toward leveraging diverse perspectives and skills, coupled with the emphasis on rapid prototyping within competitive markets, marks a departure from centralized expertise to a more global collaboration model. Despite recognizing the role of such collaborative efforts in ensuring effective governance, particularly highlighted during responses to crises such as the COVID-19 pandemic [15], the dynamics of decentralized KS demand further exploration.

Knowledge is a cornerstone of organizational success and economic value creation [6], [16], thriving through effective individual interactions within organizations [17]. The importance of KS for enhancing innovation and work efficiency is well-documented, playing a crucial role across both corporate and educational sociotechnical systems [18], [19], [20], [21], [22], [23]. Co-creation teams are seen as dynamic ecosystems gathered by the common objectives of knowledge, product, or service creation. These teams enable multidimensional interactions that drive innovation by merging diverse knowledge bases [24], [25]. Such teams can be studied to investigate the evolving nature of KS dynamics in adapting to new technological landscapes.

The increasing prevalence of online KS activities amplified by social networking platforms [26], and coupled with the transformative impact of modern technology on corporate culture [27], illustrates the interplay between the advancement of knowledge management practices and technology adoption. This synergy is vital for sustaining the effectiveness of organizations [28]. The distinction between offline and online environments for KS is crucial in understanding the dynamics of traditional versus digital collaboration systems [29]. This paradigm shift, marked by an increasing reliance on digital decentralized collaboration, is reshaping team interactions, and requires organizations to reevaluate virtual team dynamics, and leadership roles to adapt their strategies and policies to optimize the benefits of digital KS.

In virtual teams, leaders play a pivotal role in fostering effective communication and managing team challenges [10], [30], [31]. Peer production teams, exemplified by platforms such as Wikipedia [4], [32], [33], [34], [35], demonstrate a trend toward emergent, shared leadership, often without formal structure or compensation. This trend has been widely adopted [2], [36], [37] (or aimed to be adopted) in formal organizations for open innovation projects for years, which leads to a transformed leadership in team environments to more decentralized models. Larson and DeChurch highlight how these changes necessitate a reevaluation of leadership dynamics in digital platforms in a comprehensive review study [4]. The shift toward decentralized, collective leadership is supported by a body of research [1], [38], [39], [40], [41], [42], [43], [44], [45], indicating more effective leadership in digital contexts, where traditional hierarchies are less applicable and the need for agile, responsive leadership is more pronounced.

The formation and evolution of teams in digital environments significantly influence leadership dynamics [46]. Larson and DeChurch [4] highlight that in the context of digital

collaborations such as Wikipedia, team members' interactions during team formation can be instrumental in determining who emerges as a leader. Wikipedia, the peer-produced online encyclopedia, initially led by its founder Wales [47], now operates with a more decentralized leadership model; current Wikipedia page editors play a critical role in decision-making, shaping the direction [4], and flow of information on the platform. It should be noted that in this study, the leadership among content editors is referred to as "decentralized". However, we must also acknowledge that these digital teams typically benefit from the oversight of centralized platform governors and policymakers [48].

Understanding these transformed dynamics is essential for organizations operating in digital spaces, where traditional leadership models may be less appropriate. Following our digital platform example, Wikipedia, and considering the significant influence that editors wield in shaping its content and structure [4], it becomes pertinent to explore the extensive research focusing on decentralized collaborations. Numerous studies have concentrated on the collaborative dynamics of Wikipedia editors. For instance, [49] and [50] delve into the social network models within Wikipedia, exploring how they influence editorial authority and collaboration. Similarly, [51], [52], [53] provide insights into interaction patterns, signed networks of editor attitudes, and the dynamics of conflicts in Wikipedia. It should be noted that scholarly attention on Wikipedia extends beyond editor collaborations. For example, research on knowledge growth, such as the studies by [54] and [55], investigates how information is sought, processed, and expanded within the platform. Additionally, the evaluation of content quality is another area of research, with significant contributions from [56] to [57], who apply machine learning models to assess the quality of Wikipedia articles. Furthermore, comprehensive reviews of Wikipedia research, such as the recent work by [58], provide a valuable overview of the platform's multifaceted nature, encompassing editor behaviors, collaboration processes, and content development. The focus of this research, however, remains centered on the lens of Wikipedia editors as collaborators in the co-creation process to contribute to the broader understanding of collaborations in digital knowledge platforms.

In today's digital age, understanding the dynamics of information flow is crucial for organizational success [59], [60]. Employing network analysis, this study explores how team dynamics influence aspects such as information leadership, diffusion of new ideas, cooperation behaviors, and the formation of subcommunities, grounded in network science principles [61], [62], [63], [64], [65]. By evaluating the network structure across different topics and stages of artifact development, we aim to uncover insights into digital collaboration's complexities [3], [66], [67], [68].

Our study explores decentralized team dynamics over time without any central intervention, unraveling the patterns within these self-organizing systems. While this approach emphasizes the natural evolution of team interactions, there exist studies enabling central management of team collaborations, such as role assignment for avoiding conflicts [69], [70], [71]. Leveraging the findings of this study can pave the way for further research

TABLE I  
METHODOLOGICAL COMPARISON OF DATA-BASED NETWORK MODELING ARTICLES IN THE CONTEXT OF KNOWLEDGE CREATION

Article	Network Nodes	Network Edges	Network Type	Temporal Analysis Approach	Analysis Scope	Targeted Network Measures
[49]	Wikipedia editors	Common articles worked on	Common interest	Snapshot data	Non-comparative	Decentralization
[50]	Wikipedia editors	Content deleted, undeleted, restored	Direct content actions	Snapshot data	Non-comparative	Bipolarity, balance, positivity
[51]	Wikipedia editors	Direct replies in talk pages	Direct replies	Snapshot data	Topic comparison	Size and depth of discussions
[52]	Wikipedia editors	Content inserted, deleted, replaced, and restored; community elections	Signed and direct content actions	Snapshot data	Non-comparative	Ratio of the direction and sentiment-related links
[53]	Wikipedia editors	Direct replies in talk pages (opposing, agreeing, neutral)	Direct replies	Snapshot data	Non-comparative	Conflicts and editorial wars
[73]	University members	Emails	Direct replies	Dynamic evolution (over a year)	Non-comparative	Homophily emergence
[74]	Employees in the US Microsoft	Email, IM, meetings, calls	Direct replies, meetings or calls	Dynamic evolution (over two years)	Timely comparison (pre- and post-COVID)	Types of ties, subcommunity, centralization etc.
[72]	Employees in 63 U.S. firms	Emails	Direct replies	Snapshot data	Non-comparative	Interconnectivity, info speed, subcommunity, decentralization, etc.
This study	Wikipedia editors	Alternating timed interactions and content similarity	Indirect co-creation	Dynamic evolution (over 14 years)	Between-study comparison (with a hierarchical setting)	Interconnectivity, info speed, subcommunity, decentralization

on identifying circumstances under which centralized management and interventions are beneficial for higher coordination.

Building on this exploration of decentralized dynamics, this study proposes a novel framework to analyze digital co-creation, focusing on the high-level, indirect interactions that previous models overlook. Unlike approaches that concentrate on explicit direct exchanges [72], [73], [74], our method captures the essence of collaboration through subtle, interactive sequential interactions, building upon others' work over time, that form the backbone of virtual co-creation networks. By introducing an alternating timed interaction (ATI) metric, we measure the immediacy of collaboration within specific time windows, highlighting the importance of temporal proximity. Additionally, our methodology includes a content processing strategy to assess interaction depth by analyzing thematic overlaps in contributors' revisions, offering a comprehensive view of digital co-creation dynamics for different categories of artifacts. This methodology not only illuminates the complex interaction patterns within Wikipedia's editorial network but also serves as a model extendable to other digital co-creation contexts, paving the way for future research and practical applications by enhancing our understanding of digital collaboration's multifaceted nature.

Despite the evident shift toward digital platforms in sociotechnical systems, there remains a significant research gap in fully understanding the complexities of team dynamics in digital collaborative environments. This study aims to bridge this gap by conducting a comparative analysis of "Digital Decentralized Platform-Driven Collaboration" and "Traditional/Hierarchical Organizational Collaboration." This comparison aims to dissect the interaction networks—whether formal or informal—within these contrasting collaboration models.

The post-COVID era of online/hybrid communication provided the opportunity to examine this comparative study; we have utilized some recent findings on the structural measures evolution of collaboration networks in 65 U.S. organizations [72]. On the other hand, using Wikipedia as the case study for a digital platform setting, we explore the intricacies of decentralized platform-driven collaboration, analyzing its

co-creation networks to reveal the evolutionary dynamics of team collaboration [75]. Wikipedia facilitates this study with its rich data on diverse topics of articles as artifacts, with different maturity levels, and various collaboration team scales.

To better clarify the innovations and contributions of this article in comparison to existing data-based network modeling and analysis studies in the context of knowledge creation, we present Table I.

By situating our study within this context, we highlight several unique features that distinguish our research: Unlike many studies that focus solely on direct interactions such as emails and direct replies, our study incorporates indirect high-level co-creation measures, providing a novel perspective on how indirect interactions contribute to knowledge creation by capturing both ATI and content-based connections. While much of the existing literature examines either decentralized or hierarchical settings, this study uniquely compares both these settings. Our dynamic evolution analysis over 14 years (2010–2023) stands out from snapshot studies, allowing for a deeper exploration of how network structures evolve over an extended period. By addressing these aspects, our research contributes significant insights into the structural dynamics of online collaboration and informs the development of more effective digital platforms and collaborative tools.

To encapsulate our exploration focus, we anchor the study around three main research questions designed to deepen our understanding of digital collaborative dynamics.

- 1) Do collaborators behave differently when co-creating in different contexts of problems?
- 2) How are the dynamics of team structures (such as decentralized leadership, information accessibility, diffusion speed of ideas, and the formation of subcommunities and polarization) evolving over scale and time?
- 3) How different are the structural patterns of teams within a digital collaboration environment from those in traditional organizational collaborations?

The subsequent sections of this article are organized as follows. We begin with a detailed explanation of our methodology, followed by its application to the study of editorial



collaborations on Wikipedia. Then, we present an analysis of our findings and address the research questions posed earlier. The article concludes with a discussion of our study, its implications for policymakers, its limitations, and proposing some directions for future research.

### III. METHODOLOGY

In line with our exploration of digital collaboration dynamics for a comparative analysis of them with traditional organizational structures, we draw inspiration from recent studies that have performed extensive data analysis of organizational collaboration in hybrid and online environments. One such study meticulously done by Jacobs and Watts involves an exploratory analysis of a unique dataset comprising 1.8 billion messages sent by 1.4 million users from 65 publicly traded U.S. firms across various sizes and industrial sectors, investigating collaboration dynamics in contemporary settings [72]. This precedent of analyzing large-scale, real-world datasets in organizational contexts provides a robust benchmark for our investigation. Our approach initially focuses on conducting a detailed network analysis of a purely digital platform. By examining the interactions to co-create within this digital environment, we aim to derive insights that can be systematically compared with the patterns observed in traditional organizational collaborations.

An overarching view of the methodology flow is presented in Fig. 2, visually summarizing the steps from domain selection and data acquisition to network formation and calculating network characteristics. This big-picture representation aids in comprehending the comprehensive process undertaken in this study. This figure highlights the different combinations and topologies of teams working on various versions of a single artifact, illustrating the generalizability of our method in constructing all phases of collaboration evolution across diverse user topologies. The authors' primary objective is to explore how these team structures evolve over time. By tracking changes in team configurations and interactions over successive time windows, our methodology allows for an in-depth analysis of the dynamics of collaboration, the emergence of leadership, and the formation of subcommunities. This approach not only reveals the static properties of collaboration networks but also uncovers the temporal evolution and the underlying factors driving these changes. The methodology's flexibility ensures it can be applied to different domains and contexts, making it a robust tool for studying collaborative behaviors in digital environments. Detailed explanations of each step in the method diagram are provided in the remainder of this section, offering a comprehensive understanding of our analytical framework.

#### A. Domain Selection of the Artifact

In aligning our methodology with the multifaceted nature of digital co-creation platforms, this study strategically selects data across five broad categories, reflecting the insights of Gibbs, Sivunen, and Boyraz [76] regarding the significant influence of task nature in virtual teams. These categories include.

- 1) Politics, capturing discussions and collaborations surrounding key individuals in the political arena such as

“Ron DeSantis” and “Alexandria Ocasio-Cortez” and political events such as “Impeachment inquiry against Donald Trump.”

- 2) International Conflicts such as “Hong Kong protests” and “Russo-Ukrainian War” encompassing topics related to global conflicts to reflect the complexity of collaborative content in areas of international tension.
- 3) Natural and Man-Made Disasters, which includes a range of disaster-related topics highlighting collaborative efforts during global crises such as “COVID-19,” “2018 California wildfires,” and “2020 Beirut explosion.”
- 4) Technology and Business with articles such as “Tesla Model 3” and “Cryptocurrency,” focusing on the rapidly evolving sectors that significantly influence digital collaboration patterns.
- 5) Entertainment including “Parasite (2019 film)” and “Kylian Mbappé,” covering a spectrum of topics to demonstrate the diversity of interest and engagement in collaborative platforms. A full list of the sample articles used in this article exists in the first section of the Appendix document.

While this selection is tailored to our study's context using Wikipedia, it should be adapted and specialized based on the application and nature of tasks in other digital platforms within each organization. This approach ensures a balanced analysis suitable for broader applicability and adaptability in future digital co-creation explorations.

*Hypothesis 1: Collaborators behave differently when co-creating in different contexts of artifacts. In other words, the nature of the problem impacts the network formation and evolution in digital platforms.* In proposing this hypothesis, we aim to investigate whether the specific context of an artifact—be it politics, international conflicts, disasters, technology, or entertainment—has a discernible impact on the structuring and evolution of collaborative networks. This investigation acknowledges the potential for significant variability in how collaborative efforts manifest across different contexts within organizations.

Additionally, we deliberately select articles predominantly post-2010. We avoided older articles to minimize bias that could arise from Wikipedia's early developmental stages. This ensures that our analysis is reflective of more mature and established collaborative patterns on the platform. Fig. 1 is a visual representation of our domain selection, showcasing the distribution of the sample articles across each category and highlighting the creation dates of these articles.

#### B. Data Acquisition

A detailed investigation of team dynamics and information flow in digital co-creation environments is facilitated by the online traceability of collaboration steps, contrasting with the opaque nature of in-person teams, where most of the details of collaboration may remain unrecorded. In organizational contexts, while data might be anonymous to protect privacy, essential elements such as unique identifiers for each collaborator, time logs of contributions, and content topics of each

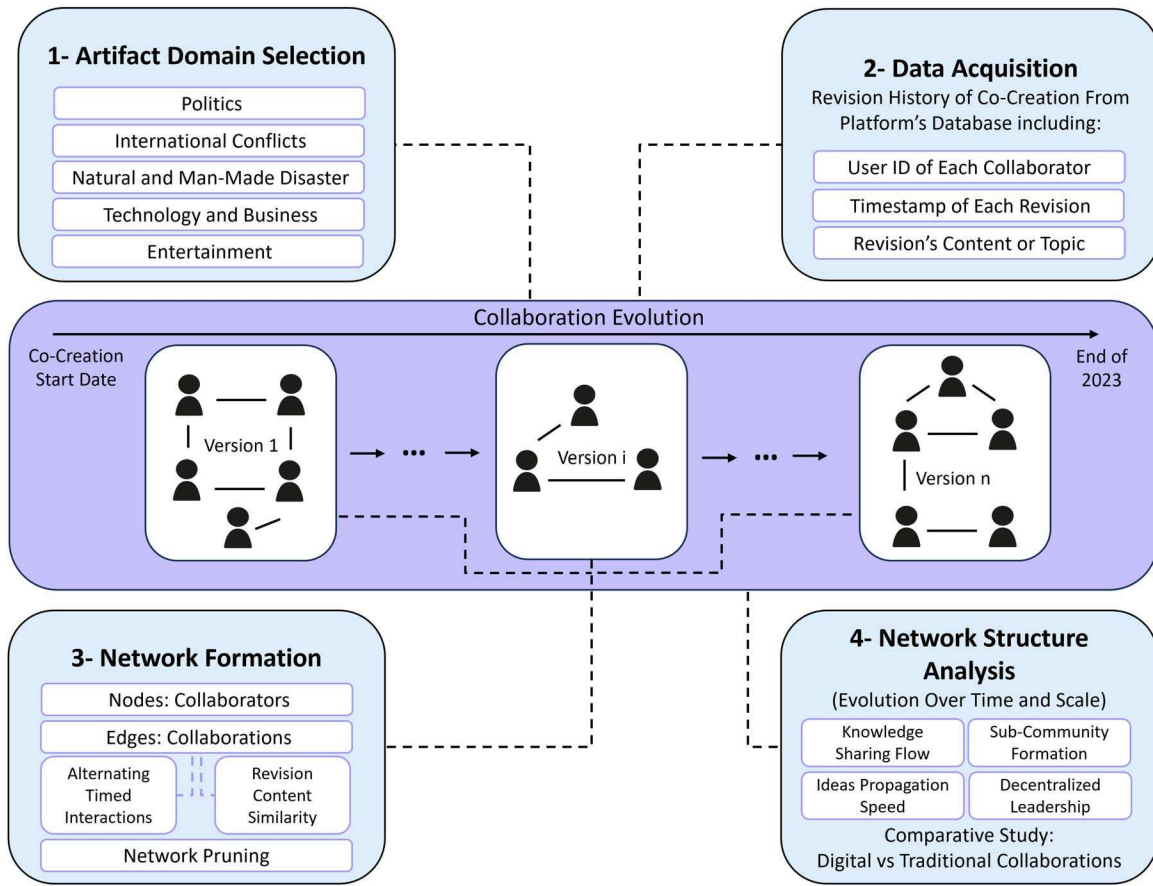


Fig. 1. Methodology overview: this figure illustrates the methodology for analyzing the evolution of collaboration within digital platforms. Different users, in various structural arrangements, work on different revisions of a problem, allowing us to observe how team dynamics evolve. Methodology consists of four main steps: artifact domain selection and scope, data acquisition, network construction model, and network analysis lenses. This approach enables a comprehensive examination of how collaborative structures form and change over time and scale, providing insights into the unique dynamics of digital co-creation.

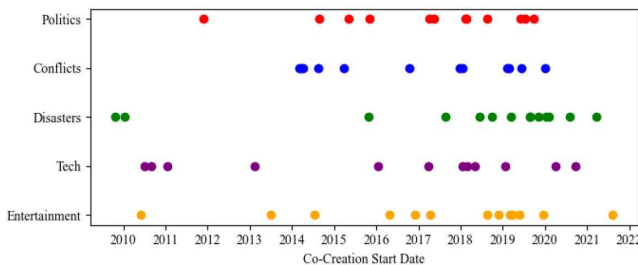


Fig. 2. Temporal distribution and the number of sample topics selected for each category. This figure facilitates an understanding of the temporal spread and categorical diversity of the sample artifacts studied. As can be seen, several sample articles with a wide range of maturity within around 13 years of evolution have been selected for this study.

task are generally accessible. This ensures a comprehensive understanding of how virtual teams operate, evolve, and lead the co-creation process. However, data acquisition, while straightforward in principle, requires meticulous processing. It involves not just the collection but also the careful parsing and interpretation of digital interactions to accurately reflect the dynamics of collaboration and co-creation.

The data acquisition phase for our study involved collecting revision histories from selected Wikipedia articles. This process was facilitated through the Wikipedia API; we captured each editor's unique identifier to map individual contributions and participation patterns within the editorial network. Additionally, timestamps of each edit were recorded to analyze the temporal dynamics of collaboration and observe how interactions evolve. Last, detailed information about the content of each revision was collected.

### C. Network Formation of Collaborators

To decipher the dynamics within a digital platform, our study employs a methodical approach to forming networks among collaborators. For each selected artifact, we establish a specific time window, measured in months, to generate a network of interactions. Notably, the commencement of this timeframe is not defined as the artifact's creation date but rather as the onset of co-creation activities; this initiation point varies, ranging from immediate postcreation engagement to several years of latency. Such differentiation ensures that our analysis starts from a period of active co-creation, yielding a more accurate

depiction of the collaboration dynamics. We define the initial parameters as follows.

- 1)  $A$ : the set of all sample artifacts under study  $a_1, a_2, \dots, a_n$ .
- 2)  $P_a$ : the set of periods (timeframes) for each artifact  $a$ .
- 3)  $U_{a,p}$ : the set of users that have edited the artifact  $a$  in a specific period  $p$ .
- 4)  $R_{a,p}$ : the set of revisions  $r$  in a specific timeframe  $p$  on artifact  $a$ .

Network graphs of collaborators for each article within the predefined timeframe are constructed as follows: nodes are represented by the unique identifiers for users in  $U_{a,p}$ . Edges are constructed based on two principal dimensions—temporal and content interactions among editors.

1) *Temporal Dimension*: An interaction time threshold (e.g., a few hours/days) is set as a meta-parameter of our model. Within this parameter, reciprocal interaction activities between two collaborators are quantified. To achieve this, we introduce a measure of mutual interaction, accounting for indirect interactions, which are evidenced by users building upon each other's edits or contributions. Addressing the challenge of inferring indirect interactions, we will use the notion of ATIs. This pattern, characterized by alternating activities between individuals (e.g.,  $i \rightarrow j \rightarrow i \rightarrow j \rightarrow i \rightarrow j$ ) within the specified time window, serves as an indirect indicator of active and reciprocal engagement. It stands in contrast to nonalternating patterns (e.g.,  $i \rightarrow i \rightarrow i \rightarrow j \rightarrow j \rightarrow j$ ). Specifically, if a sequence of edits by “i” and “j” occurs close enough (within the defined time threshold) in the artifact, this interaction accrues a temporal weight of 1 in the connection between these two editors. This weight increases proportionally with the frequency of such sequential interactions within the defined threshold. The importance of the timely closeness of two users' activities in capturing their interaction lies in the fact that online co-creation artifacts evolve over time and are updated—either slightly or significantly—each time a user makes a change or adds/eliminates parts. Consequently, the artifact might look completely updated after a few days or months, depending on the artifact's engagement rate.

2) *Content Dimension*: To quantify the content dimension of interactions among Wikipedia editors, we analyze the thematic similarities between their revisions. This involves identifying the main topics of an article's subsections and employing natural language processing (NLP) techniques to preprocess and analyze the text of revisions for thematic content. By preprocessing the text—tokenizing, removing stopwords, and filtering nonalphanumeric tokens—we distill the revisions to their informative essence. We then match the processed text to article topics using a topic-allocation strategy, assigning scores based on the occurrence of topic-related keywords within the revisions. For each pair of editors  $i$  and  $j$  working on the same artifact, we determine the overlap in topics they have contributed to by comparing the sets of topics associated with their revisions. The intersection of these sets indicates their thematic alignment. The strength of the connection between two editors is measured by the size of this intersection, normalized by the total number of subsections in the article to account for variations in topic breadth. This normalized weight represents the content-based

connection strength between editor pairs, offering insights into the thematic coherence of their collaborations.

Given the definitions, the total weight of the connection between collaborators  $i$  and  $j$ , denoted by  $W^{ij}$ , is calculated as the product of the temporal dimension  $W_T^{ij}$  and the content dimension  $W_C^{ij}$  of these weights

$$W^{ij} = W_T^{ij} \cdot W_C^{ij} \quad (1)$$

where

$$W_T^{ij} = \sum_{k=1}^{|\tau|-1} I(u_k \neq u_{k+1}, |\tau_{k+1} - \tau_k| \leq \bar{t}). \quad (2)$$

Let  $\tau$  be the set of timestamps of edits for both collaborators and  $\bar{t}$  be the interaction time threshold. Then  $\tau_k$  and  $\tau_{k+1}$  are consecutive timestamps in  $\tau$ , and  $u_k$  and  $u_{k+1}$  are the corresponding users for these timestamps.  $I(\cdot)$  is an indicator function that is one if its argument is true and 0 otherwise.

For the content weights, if we denote the set of sections of each artifact  $a$  as  $S_a$ , and the subsets of these sections that each collaborator  $i$  has worked on as  $S_i^a$ , then the content weight  $W_C^{ij}$  for editors  $i$  and  $j$  in artifact  $a$  is the proportion of overlapping sections they worked on, given by the following equation:

$$W_C^{ij} = \frac{|S_i^a \cap S_j^a|}{|S_a|}. \quad (3)$$

To create an interaction network for the specific topic in each time window, the history of all the revisions (timely ordered) done on that article needs to be retrieved. Then  $W_T$  and  $W_C$  for all the user pairs active in that period are calculated.

Then the interaction network for that timeframe is created by enabling an edge  $E_a^{ij}$  between all the pairs of collaborators ( $i$  and  $j$ ) that have had interactions based on nonzero  $W^{ij}$  value. To ensure the analysis focuses on meaningful interactions within the digital platform, the study employs a crucial step of network pruning

$$E_{a,p} = ij | W_{a,p}^{ij} \geq \bar{w}. \quad (4)$$

This process involves setting a threshold for the minimum weight,  $\bar{w}$ , required for an edge to be considered significant. By filtering out less impactful collaboration edges, the process distills the network to its most influential connections, offering a clearer view of the core collaborative interactions within each digital artifact. In the end, let  $G_{a,p}$  be the network graph for artifact  $a$  and the time window  $p$  with all the remaining edges after pruning

$$G_{a,p} = (U_{a,p}, E_{a,p}). \quad (5)$$

The giant connected component (GCC) of the constructed graph is kept for the rest of the analysis. All these steps are summarized in Algorithm 1. The algorithm complexity of constructing the interaction network for one artifact within a single period is primarily driven by the number of users and the number of revisions. Specifically, the overall time complexity is  $O(|U_{a,p}|^2 |R_{a,p}|)$ , where  $|U_{a,p}|$  is the number of users and  $|R_{a,p}|$  is the number of revisions within the timeframe. This complexity encompasses the steps of calculating temporal and content-based weights and constructing and pruning the network graph,

**Algorithm: 1-** Co-Creation Network Construction

---

```

1: for artifact  $a \in A$  do
2:   Determine  $P_a$ .
3:   for period  $p \in P_a$  do
4:     Retrieve sorted revision history  $R_{a,p}$ .
5:     Initialize  $W_{a,p} = \emptyset$ .
6:     for each user pair  $ij$  do
7:        $R_{a,p}^{ij} = R_{a,p}^i \cup R_{a,p}^j$ 
8:       Initialize  $W_T^{ij}, W_C^{ij} \leftarrow 0$ .
9:       for  $(r_k, r_{k+1}) \in R_{a,p}^{ij}$  do
10:         $u_k, \tau_k \leftarrow$  user, timestamp of  $r_k$ 
11:         $u_{k+1}, \tau_{k+1} \leftarrow$  user, timestamp of  $r_{k+1}$ 
12:        if  $u_k \neq u_{k+1}$  and  $\tau_{k+1} - \tau_k \leq \bar{t}$  then
13:           $W_T^{ij} \leftarrow W_T^{ij} + 1$ .
14:          Calculate  $W_C^{ij}$  (Equation 3).
15:        end if
16:      end for
17:      Calculate  $W_{a,p}^{ij}$  (Equation 1).
18:      Store  $(ij, W_{a,p}^{ij})$  in  $W_{a,p}$ .
19:    end for
20:  end for
21: end for
22: for  $a$  and  $p$  do
23:   Initialize  $G_{a,p} = \emptyset$ .
24:   for  $ij$  in  $W_{a,p}$  do
25:      $G_{a,p} \leftarrow G_{a,p} \cup (ij, W_{a,p}^{ij})$ .
26:   end for
27:   Calculate  $\bar{w}$  based on  $W_{a,p}$ .
28:   Prune  $G_{a,p}$  based on  $\bar{w}$  (Equation 4).
29:   Store  $G_{a,p}$  for structure analysis.
30: end for

```

---

and the dominant action in all three steps is going through all pairs of users, and for each pair, processing the revisions.

#### D. Network Structure Analysis

In the next step, we employ statistical network analysis techniques. While there are plenty of network statistics available, in our approach, we have restricted our statistical network analysis to a selected set of network measures, carefully chosen not only for their ability to capture essential aspects of collaborative behavior but also to allow for a balanced comparison with existing studies on traditional hierarchical organizational structures.

To understand the extent and nature of collaboration, we focused on measuring team cohesion by calculating the average number of interactions each member has within the network. This aspect is crucial as it sheds light on the potential for knowledge transfer and potential innovation [67], [77], [78], [79]. Another dimension we examined was assessed by determining the average number of steps required to connect any two members of the network, providing insights into the network's capability to facilitate the rapid spread of information [61], [63], [80], [81]. Additionally, we explored the tendency of network members to form closely knit groups. This measure is indicative of the level of collaboration and open communication within the network,

highlighting the potential for collaborative efforts to be concentrated among certain groups [67], [68], [82], [83]. Finally, we looked at the network's structural propensity to centralize information flow. By assessing which members predominantly control or influence the flow of information, we gained insights into the network's architecture regarding how it enables certain individuals to act as crucial connectors, impacting the overall communication process [84], [85], [86], [87], [88].

In line with this analytical framework, we propose our next hypothesis as follows.

*Hypothesis 2: The characteristics of team structures within collaborative networks exhibit discernible patterns that evolve over scale (size) and time (age).* This hypothesis aims to explore the dynamic interplay between team structures and the evolving nature of digital collaboration.

Following the insights provided by Jacobs and Watts in their comprehensive 2021 study, "A large-scale comparative study of informal social networks in firms," we find a robust foundation for our analysis. Jacobs and Watts' pioneering work, which involved a detailed analysis of anatomized email data across U.S. firms, revealed notable variations in network metrics influenced by organizational size. Their study highlighted the interplay between microlevel network structures and overarching organizational properties, even though clear correlations with firm age, and the industry that the firms work in were not established [72].

The strategic alignment of network analysis with existing literature enables us to draw a parallel and conduct a meaningful comparison between the network patterns observed in digital platform-driven collaboration and those prevalent in traditional organizational settings. In this regard, we introduce our final hypothesis as follows.

*Hypothesis 3: The structural patterns of networks within digital collaboration environments, such as Wikipedia, will exhibit significant differences when compared to those found in traditional organizational collaborations.* This hypothesis stems from the premise that the inherent characteristics of digital platforms—such as their decentralized nature, the scalability of interactions, and the digital traceability of collaboration processes—cultivate distinctive network structures. These structures are likely not to align with those formed within the more bounded and hierarchical contexts of traditional organizations. By investigating this hypothesis, we aim to shed light on the unique dynamics of digital collaborative networks and how they diverge from conventional organizational network patterns, thereby contributing to a deeper understanding of digital co-creation processes.

## IV. IMPLEMENTATION AND RESULTS

For this study, we constructed interaction networks for Wikipedia collaborators at 6-month intervals, starting from the collaboration onset of each article until the end of 2023. Consequently, depending on when each sample artifact was created, we generated multiple collaboration networks for every article, each representing a distinct 6-month period. After this, the ATI and content-based weights ( $W_T^{ij}$  and  $W_C^{ij}$  with the averages of



1.15 and 0.72, respectively, for the interaction time threshold of 48 h) for all the 29 149 pairs existing in all the sample artifacts are calculated, and the 6-month period collaboration networks are generated based on them. More detailed information regarding these weights exists in the second section of the Appendix document.

Following the construction of these initial networks, we refined our analysis through a process of network pruning. By setting a threshold for the minimum weight required for an interaction to be considered significant at 70%, we focused on the most impactful collaborative interactions. Although there is no specific optimal weight threshold for all the collaboration networks, we tested several values and selected 70% as it was sufficient to keep the most meaningful interactions in this study. This threshold was determined by comparing the original graphs' edges with their GCC and choosing the value where they overlapped sufficiently (more than 0.95), indicating efficient network construction and pruning. Our study concentrates on analyzing the GCC of these networks, which represents the primary structure of collaborative interactions within Wikipedia.

An additional step of our methodology involved filtering preactive time windows from the network data: we considered a network as significant when its size exceeded 15% of its maximum size throughout its entire lifetime for that article. There exist some samples for this time-filtering, the reasoning behind choosing this threshold for our dataset, and also the statistical summary of the data before this filtering in the Appendix document. In addition to these steps, we also assessed the maturity of each network by calculating the age of the artifacts in months, from the start of collaboration to the onset of each network formation.

After filtering the preactive time windows out, our refined dataset comprised 789 out of the original 839 networks, spanning various topics such as politics, conflicts, disasters, technology, and entertainment (originally included 168, 155, 145, 192, and 179, respectively). This consolidation resulted in a dataset forming the basis for our subsequent analysis. A summary of these networks, including data such as the unique number of topics, network counts, and the range of nodes in the GCCs for each category, is presented in Table II. This table provides a snapshot of our dataset's breadth and diversity, offering insights into the scale and temporal spread of the articles analyzed.

The dataset of collaboration networks also reveals that, on average, they comprise ten nodes. The average degree (D) stands at 4, indicating a moderate level of interconnectedness among editors. With a clustering coefficient (C) of 55%, the networks tend to form tight-knit groups. The networks exhibit an average of 1.6 for the average shortest path length (L), facilitating efficient information flow, while the average centralization (BC) is less than 1%, suggesting a decentralized network structure. The age of these networks spans a broad range, from as recent as 0 months to as mature as 158 months. Additionally, it should be noted that the number of edges in the GCC over the number of edges in the original pruned nonconnected network, varies between 50% and 100%, with an overall average of 97%. This

TABLE II  
SUMMARY OF CONSTRUCTED NETWORKS BY CATEGORY

Category	#Topics	#Networks	Min Start Date	Max End Date	Min #Nodes	Max #Nodes
Politics	12	148	2011-11-28 00:18:34+00:00	2023-11-30 22:57:44+00:00	2	29
Conflicts	12	155	2014-03-01 23:01:38+00:00	2023-12-23 17:07:20+00:00	2	27
Disasters	14	144	2009-10-25 21:51:24+00:00	2023-12-10 21:02:50+00:00	2	27
Tech	12	176	2010-01-03 08:46:49+00:00	2023-11-02 21:03:38+00:00	2	31
Entertainment	13	166	2010-06-03 00:19:30+00:00	2023-12-18 13:54:27+00:00	2	36

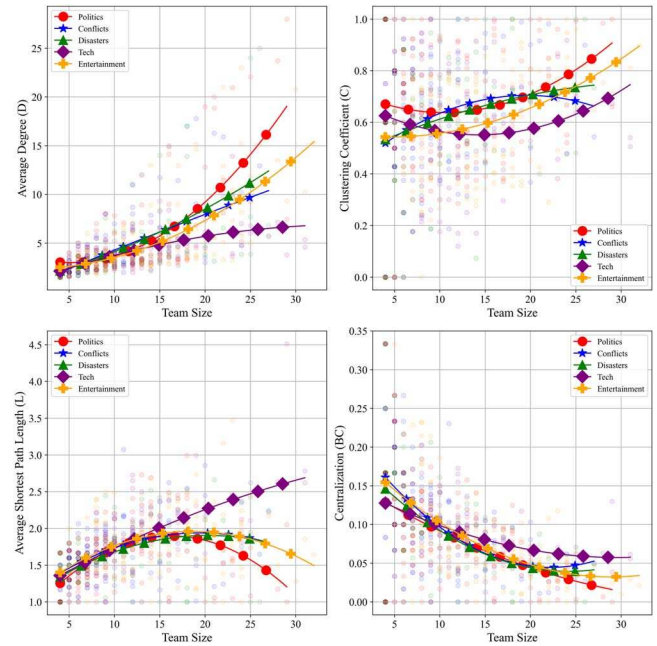


Fig. 3. Network structure patterns when the teams grow in size per artifacts category: the second-order regression lines are fitted on the characteristics of all the network graphs generated and kept for analysis. Each collaboration network is defined per 6-month time window from the onset of the meaningful interactions until the end of 2023.

high ratio underscores the robustness of the GCC in our analysis. For a detailed summary of statistics of the results and a visual presentation of the distribution of these characteristics, please refer to the Appendix document.

It should also be noted that networks with fewer than four nodes are excluded from all our subsequent visualizations and analysis. This decision was based on the rationale that very small networks might not accurately represent the collaborative patterns we sought to study.

#### A. Network Dynamics and Size of the Team

As the starting point of our exploration, Fig. 3 plots the pattern lines fitted on various team structure characteristics against the increasing team size for all the remaining graph networks constructed. The figure comprises subplots that illustrate



TABLE III  
REGRESSION ANALYSIS OF NETWORK STRUCTURE METRICS  
AND TEAM SIZE

Model Variables	Category	Size p-value	Size <sup>2</sup> P-value	Prob (F-statistics)	Adjusted R <sup>2</sup>
Average Degree vs Team Size	Politics	0.672	≤ 0.001*	≤ 0.001*	0.548
	Conflicts	≤ 0.001*	0.768	≤ 0.001*	0.480
	Disasters	0.016*	0.430	≤ 0.001*	0.483
	Tech	≤ 0.001*	0.005*	≤ 0.001*	0.490
	Entertainment	0.027	≤ 0.001*	≤ 0.001*	0.571
Clustering vs Team Size	Politics	≤ 0.001*	0.005*	≤ 0.001*	0.191
	Conflicts	≤ 0.001*	0.001*	≤ 0.001*	0.163
	Disasters	≤ 0.001*	0.002*	≤ 0.001*	0.187
	Tech	≤ 0.001*	0.016*	≤ 0.001*	0.084
	Entertainment	0.874	0.293	≤ 0.001*	0.109
Average Shortest Path vs Team Size	Politics	≤ 0.001*	≤ 0.001*	≤ 0.001*	0.382
	Conflicts	≤ 0.001*	≤ 0.001*	≤ 0.001*	0.389
	Disasters	≤ 0.001*	≤ 0.001*	≤ 0.001*	0.511
	Tech	≤ 0.001*	0.003*	≤ 0.001*	0.600
	Entertainment	≤ 0.001*	≤ 0.001*	≤ 0.001*	0.217
Centralization vs Team Size	Politics	0.349	0.047*	≤ 0.001*	0.086
	Conflicts	0.303	0.069	0.011*	0.045
	Disasters	0.018*	0.008*	0.023*	0.040
	Tech	0.427	0.178	0.120	0.013
	Entertainment	0.015*	0.561	≤ 0.001*	0.244

Note: \*Significant at the 0.05 level.

the relationship between team size (number of nodes of each constructed GCC) and four key network characteristics: each is represented in its subplot, where categorized and color-coded based on five topic categories: politics (red), conflicts (blue), disasters (green), tech (purple), and entertainment (orange).

For each topic category, a second-order polynomial regression line is fitted to the data points, showing the trend of each characteristic as the team size increases. These regression lines help to visualize the general pattern and commonalities among different topic categories, providing insights into how team size affects the structural properties of the networks. The subplots are labeled with the team size (x-axis) and the respective network characteristic (y-axis) to facilitate interpretation. The fitted polynomial lines reveal the underlying trends and variations in team structure characteristics.

This exploration is enriched by the regression model results presented in Table III, which examines the relationship between the team size and each of its structure characteristics. Quadratic terms were added to the regression model to capture nonlinear relationships. The independent variables included team size (number of nodes in GCC) and its squared term size<sup>2</sup>, while the dependent variables were the network characteristics. For each category (politics, conflicts, disasters, tech, entertainment), we conducted separate regressions to analyze the effects within specific contexts. The regression results were summarized in terms of p-values for both the linear and quadratic terms, coefficients, F-statistic probabilities, and adjusted R-squared values, providing a comprehensive view of the model fit and the significance of the predictors. Values with p-values less than 0.05 are presented with a \* to indicate significant relationships.

The regression analysis examining the relationship between network measures and the size of the teams revealed insightful patterns. Notably, in some categories, we observed a lack of significant relationships between some of the network characteristics and the team size, underscoring the influence of content specificity on our results. However, most categories displayed

significant associations, highlighting how the nature of each artifact shapes collaborative dynamics. Across all categories and characteristics, our analysis unveiled a spectrum of relationships with team size. Predominantly, these relationships are nonlinear, as evidenced by the significance of both linear and squared terms in the regression models. The degree of explanation, represented by R-squared values, varied across characteristics and categories, indicating that some models better fit the data than others. A visualization of these trends including the scatter plots of the team's network data points exists in the Appendix document. The findings of this section are summarized in Table IV which is explained in detail in what follows.

1) *Team Cohesion and Interconnectivity Versus Team Size:* When examining the relationship between interconnectivity captured by the average degree of the team and the number of members collaborating, two perspectives emerge from existing literature. First, Jacobs and Watts' Perspective posits that the average degree does not vary with size which is validated by their massive organizational data analysis; it is based on the assumption that while individual cognitive capacities may vary, the overall average should remain relatively stable across organizations [72]. Second, an alternative viewpoint suggests that the average degree should increase with organizational size due to greater availability of potential contacts [89] or flatter managerial structures [90].

Our findings align with the second viewpoint for mainly two reasons pertinent to digital platforms.

- Their artifact-centered design:* Its digital platforms are structured to focus on specific artifacts, allowing individuals to concentrate on one problem at a time. This design potentially enhances cognitive capabilities during interactions, as attention remains focused on a singular issue.
- Their decentralized leadership:* In contrast to traditional hierarchical organizational structures, digital platforms often feature decentralized leadership. This autonomy allows for parallel interactions with potential collaborators, facilitating an increase in each team member's degree as more people engage with the same problem. To summarize, the connectivity of each team member is given both the opportunity and authority or permission to increase when more people are available to work on the same problem.

This increasing trend when a team grows in size in digital platforms—whether it be a significant quadratic relationship as observed in the political domain, a significant linear relationship as in conflicts and disasters, or more complex relationships in technology and entertainment domains—contrasts with traditional organizational structures. It presents an opportunity for the future of co-creation platforms, where increased interaction links could mean easier information exchange. However, caution is warranted: excessive KS might be detrimental, leading to fast convergence on solutions, especially for complex artifacts. This balance between beneficial and excessive information exchange is critical for the effective functioning of collaborative platforms [91].

2) *Subcommunity Formation and Polarization Versus Team Size:* In this section, we investigate the intricate relationship

TABLE IV  
SUMMARY OF THE COMPARATIVE ANALYSIS BETWEEN THE TEAM DYNAMICS WHEN THE TEAM GROWS IN SIZE IN DIFFERENT SETTINGS

Team Measure	Knowledge-Sharing Flow	Subcommunity Formation	Ideas Propagation Speed	Decentralized Leadership
Social network characteristic	average degree	clustering coefficient (transitivity)	average shortest path length	average centralization coefficient (opposite of it)
Previous studies	constant (cognitive limit) increase (more potential contacts)	decrease with different rates	increase	all increase, decrease, and staying constant have been observed
Digital platform setting behavior	increase	increase in general (median: 63%)	increase: up to mid-size teams decrease: for larger teams	decrease (almost context-independent max: 33%)
Possible cause of the differences	(1) opportunity enabled (2) authority given	working (agreeing/disagreeing) on similar interest is enabled	logical consequence of previous measures	(1) platform structure (2) autonomous users

between team size and the tendency of subcommunity formation, as indicated by clustering patterns, within digital collaborative platforms. This is captured by the transitivity measure. This analysis draws upon various network science models, contrasting their predictions with our empirical observations across different topic categories. Initially, models such as random graph theory [61] suggest a decrease in subcommunity clustering with increasing team size. In contrast, some small-world models [61], [92], and those assuming hierarchical structures [93] propose a constancy in clustering despite team growth. However, the scaling proposed by Jacobs and Watts offers an intermediate view, suggesting that clustering does not fit neatly into a constant or inversely proportional pattern [72]. In our analysis of digital platforms, we observed distinct trends in clustering across different topic categories, diverging from these traditional models:

- Conflict and disaster topics*: Here, we observed a consistent increase in subcommunity clustering, leveling off in very large teams. This trend is likely influenced by the time-sensitive and critical nature of these topics.
- Politics and technology categories*: A U-shaped pattern was evident, with initial dispersal in subcommunities followed by a sharp convergence into tightly knit clusters as team size increases.
- Entertainment domain*: Although an increasing trend in clustering was observed, it did not present significant patterns, as detailed in Table III, suggesting a more fluid team structure in this domain.

This observed pattern contradicts traditional network science literature, including random graph models, small-world models, and the findings of Jacobs and Watts. Significantly, the range of clustering coefficients in Wikipedia networks (with a mean of 0.62 and median of 0.63) is substantially higher than those observed in email-based interaction networks with a maximum of 0.30 [72], suggesting that digital platforms foster more clustering as teams working on a single artifact within a defined context grow larger.

The observed trend of increasing clustering in virtual teams, especially as they scale up, likely mirrors the dynamics of subcommunity formation and polarization within digital platforms; in larger virtual teams, individuals tend to form clusters around specific interests, expertise, or viewpoints. This is especially pronounced in domains such as politics, where the breadth of the subject matter can lead to specialized clusters. The U-shaped pattern observed in politics suggests that as more individuals join the team, they initially explore various aspects

of the topic independently before eventually forming tight-knit clusters around specific subtopics or approaches. On the other hand, in certain contexts, especially in conflicts and disasters, the urgency and high stakes can intensify polarization within teams. Members may cluster around specific solutions or approaches, driven by their experiences, expertise, or emotional responses to the situation. Such polarization, while potentially limiting diverse viewpoints, can also lead to a more in-depth exploration of different strategies.

In a nutshell, as virtual teams grow, the complexity and diversity of the topics lead to more pronounced subgroup formations. This has been one of the most surprising findings of this study. The authors believe that understanding this trend is crucial for organizations leveraging digital platforms for collaboration. While forming clusters can enhance focus and depth in specific areas, it also poses the risk of creating echo chambers that might suppress innovation and diverse perspectives.

3) *Information Propagation Speed Versus Team Size*: Previous studies in network science have consistently suggested that the speed of information propagation captured by the average shortest path length ( $L$ ) increases logarithmically with team size. As Newman (2003) posits, in networks where the average degree ( $D$ ) is constant,  $L$  is expected to rise logarithmically with the number of team members [93]. Jacobs and Watts further elaborate that  $L$  is highly negatively correlated with  $D$ , especially in large firms with hierarchical team structures. In such environments, despite the growth of the firm, if individuals maintain a constant number of contacts, the paths between people stay constant [72].

However, our analysis reveals a distinct pattern that departs from these conventional findings: for all the categories, we observe a complex relationship; initially, there is an increase in  $L$  for mid-size teams, followed by a decrease in  $L$  as the network size expands significantly. This pattern aligns with the logic of previous studies. Given that  $D$  increases in our data (as elaborated earlier), we anticipated a sharp decline in  $L$  toward the end, as the paths between two people should become shorter due to the higher degree of the members and the negative correlation between  $L$  and  $D$  (around 40% in our dataset). However, it should be noted that in conflicts, disasters, and technology domains, we witness a slower final decrease in  $L$ ; this observation is intriguing, particularly considering the simultaneous increase in both  $D$  and  $C$  as team size grows. While several factors could contribute to this trend, one compelling explanation aligns with our earlier hypothesis regarding polarization within these domains.

In summary, the trend we uncover in the speed of information diffusion versus the team size sheds new light on the dynamics of digital collaborative networks, highlighting deviations from established network science theories. The divergence in trends across different categories emphasizes the impact of domain-specific factors on virtual team structures.

4) *Centralized Leadership Among Subgroups Versus Total Team Size*: Jacobs and Watts' study presents an intriguing insight into organizational network structures, suggesting that centralization remains constant as firms grow, typically the value of 0.84 for betweenness centrality (BC). They also note the ambiguity in prior literature regarding the expected variation of centralization with firm size [72]. Contrastingly and interestingly, our study unveils a unique significant pattern in the realm of digital collaboration. We observe that the BC in networks formed by digital platform collaborators is not only lower than 0.33 for all the constructed networks during the entire timeline of the study for all the sample artifacts but also tends to decrease as the team size increases. It should be noted that the visual declining trend in BC is not statistically significant in the technology domain but is still very low. Decentralization is a crucial factor in the co-creation of knowledge, as stated in a network-driven study of Wikipedia editors: "The more decentralized the editing of an article, the better the article represents a consensus" [49].

Several factors could potentially explain this difference from the previous studies.

- Nature of organizations*: previous research mostly focuses on hierarchical settings where centralization often pertains to decision-making being concentrated at higher levels. In contrast, Wikipedia's editing environment represents a digital, open-source platform that inherently fosters a distributed collaboration.
- Type of collaboration*: In traditional firms, specific mechanisms may be implemented to sustain a level of centralization as the organization grows. Conversely, Wikipedia exemplifies a decentralized model for knowledge generation despite its central platform governance, where editors contribute autonomously. As such networks expand, the necessity for central nodes diminishes, leading to a decrease in BC.
- Cultural and behavioral factors*: The culture of collaboration in open-source communities, such as Wikipedia, emphasizes distributed, egalitarian participation. This cultural inclination toward decentralization contrasts with the more hierarchical and centralized control often seen in corporate environments.

In essence, our findings about the average centrality of leadership between various subgroups and team size on digital platforms such as Wikipedia challenge traditional notions of network centralization, highlighting the distinct dynamics of digital collaborative environments. This divergence not only reflects the structural and cultural differences between digital platforms and traditional organizations but also underscores the importance of considering these unique dynamics when designing and managing digital collaborative systems.

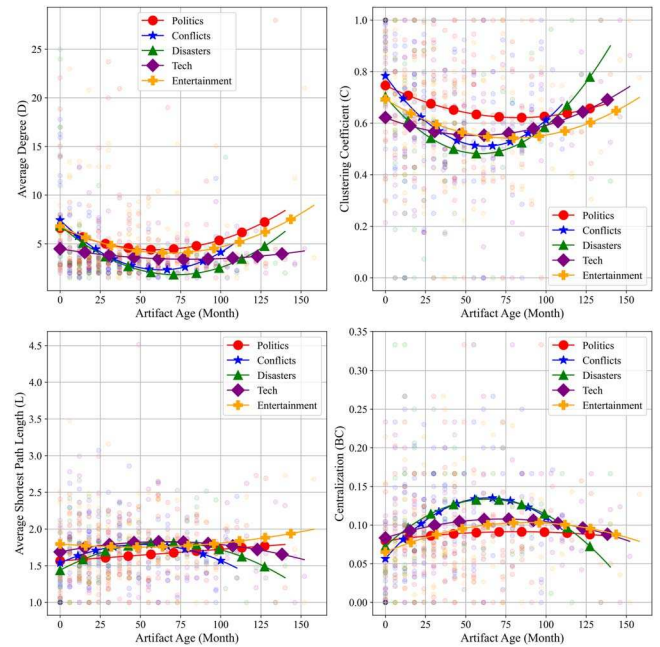


Fig. 4. Network evolution patterns when the artifacts age for each category: the second-order regression lines are fitted on the characteristics of all the network graphs generated and kept for analysis as the artifact ages.

## B. Network Dynamics and the Maturity of the Artifact

This study ventures beyond the organization's age and, instead, pivots to explore how the age of the artifacts influences the structural dynamics of the teams working on them. This exploration is grounded in the premise that the lifecycle of an artifact—its evolution from inception to maturity—plays a pivotal role in shaping the collaboration patterns and network structures surrounding it.

Similar to the method used for unrevealing patterns in the previous section, for each topic category, a second-order polynomial regression line is fitted to the data points, showing the trend of each characteristic as the artifact matures. The subplots are labeled with the artifact age on the horizontal axis and the respective network characteristic on the vertical axis to facilitate interpretation.

Our comprehensive analysis across various domains has revealed intricate relationships between the age of artifacts and multiple aspects of team dynamics. Each aspect—team cohesion, subcommunity formation, information propagation, and centralized leadership—shows unique patterns of interaction with the artifact's age, reflecting the evolving nature of collaborative efforts. Based on the detailed analysis, it becomes evident that the relationship between team structures and artifact maturity varies across different content categories as shown in Fig. 4. This section of the article presents the details of these relationships, examining the extent to which the age of an artifact influences the structure of the collaborative networks on Wikipedia. The regression models of this analysis were defined with age (artifact's age) and  $\text{age}^2$  as independent variables and each network measure as dependent variable.



TABLE V  
REGRESSION ANALYSIS OF NETWORK STRUCTURE METRICS  
AND ARTIFACT AGE

Model Variables	Category	Age p-value	Age <sup>2</sup> P-value	Prob (F-statistics)	Adjusted R <sup>2</sup>
Average Degree vs Artifact Age	Politics	0.003*	0.005*	0.013*	0.051
	Conflicts	≤ 0.001*	≤ 0.001*	≤ 0.001*	0.280
	Disasters	≤ 0.001*	0.003*	≤ 0.001*	0.124
	Tech	0.079	0.140	0.157	0.011
	Entertainment	≤ 0.001*	≤ 0.001*	≤ 0.001*	0.081
Clustering vs Artifact Age	Politics	0.039*	0.124	0.050*	0.029
	Conflicts	≤ 0.001*	0.002*	0.000*	0.111
	Disasters	0.005*	0.004*	0.014*	0.075
	Tech	0.159	0.098	0.220	0.008
	Entertainment	0.007*	0.021*	0.019*	0.037
Average Shortest Path vs Artifact Age	Politics	0.847	0.918	0.687	0.000
	Conflicts	0.017*	0.017*	0.050*	0.031
	Disasters	0.004*	0.007*	0.016*	0.072
	Tech	0.221	0.208	0.450	0.000
	Entertainment	0.248	0.176	0.376	0.000
Centralization vs Artifact Age	Politics	0.269	0.402	0.451	-0.003
	Conflicts	≤ 0.001*	0.002*	≤ 0.001*	0.115
	Disasters	0.007*	0.009*	0.026*	0.061
	Tech	0.100	0.128	0.254	0.006
	Entertainment	0.021*	0.077	0.034*	0.029

Note: \*Significant at the 0.05 level.

For each domain category and each network measure, separate regression models were run to understand the unique effect within different contexts. The model results of this section are presented in Table V. Key statistics of regression modes (p-value of the independent variables, the F-statistics probability, and the adjusted R-squared are extracted; values with p-value less than 0.05 are shown by \* to indicate significant relationships.

1) *Team Cohesion and Interconnectivity Versus Artifact's Age*: In the realm of politics, we observed a subtle yet significant correlation between the depth of collaboration and artifact age. Conversely, in the conflicts and disasters domains, a more pronounced relationship suggests that the maturity of artifacts plays a more substantial role in shaping collaborative dynamics. Across all categories, teams formed around newer articles exhibited higher interconnectivity, which tended to decrease with age, only to increase again in later stages. Notably, the technology domain stood apart, showing no meaningful relationship, likely due to the rapid evolution characteristic of this field.

2) *Subcommunity Formation and Polarization Versus Artifact's Age*: The clustering tendency's correlation with artifact age varied across domains. Except for technology, all categories showed a significant, moderately strong correlation. Typically, articles began with high clustering, which decreased over time, and then increased as the artifacts matured. This pattern, which is interestingly similar to the previous pattern, is indicative of the dynamic evolution of information and understanding over time.

3) *Information Propagation Speed Versus Artifact's Age*: In terms of information diffusion speed, captured by the average shortest path length, significant correlations with artifact age were only found in the domains of disasters and conflicts, exactly opposite to their average degree and clustering formation patterns. The statistically insignificant relationships among the

rest of the domains suggest that while an artifact's maturity might be a factor for some artifacts, it interacts with a multitude of other elements influencing how quickly ideas spread within these teams.

4) *Centralized Leadership Among Subgroups Versus Artifact's Age*: Finally, the relationship between centralized leadership within subgroups and the artifact age was significant in the disasters, conflicts, and entertainment categories showing a sharp increase and then a decrease in the centralization of the collaboration teams. However, its explanatory power was somewhat limited, indicating that while artifact age affects network centralization, it is not the only determinant.

These findings reveal an interesting dynamic between the evolution of collaborative efforts and the maturity of artifacts, particularly pronounced in the conflicts and disasters categories. In these domains, a distinct pattern emerges, characterized by a high level of initial engagement, as indicated by increased team interconnectivity and group cohesiveness, alongside efficient information propagation and centralized decision-making. This heightened activity in the early stages can be linked to the urgent, time-sensitive nature of these topics, fostering immediate and intense collaboration. As events unfold and the immediacy drops, a noticeable shift occurs. We see a decrease in both team interconnectivity and group cohesiveness, along with a lengthening of information propagation paths and a reduction in centralized decision-making. This change reflects a transition in the collaborative focus, moving away from the immediate response to a more prolonged and evolving engagement with the topic. However, as the artifacts grow older, network activity is resurgent. This is marked by a renewed increase in team interconnectivity and group cohesiveness, accompanied by a decrease in information propagation paths and centralized decision-making. Such a revival in collaborative efforts can be attributed to several factors, such as anniversary reflections, renewed public or scholarly interest, or ongoing impacts of the events. This pattern highlights the cyclical nature of collaboration in these domains, driven by evolving public and academic interest over time.

In summary, this relationship between artifact age and collaborative network dynamics sheds light on the adaptive nature of collaboration, particularly in areas marked by immediacy and long-term evolution. Understanding these patterns is crucial for managing and fostering effective collaborative environments over the lifespan of a project, especially in digital platforms where temporal dynamics play a significant role.

*The key finding regarding Hypothesis 2: team structures reveal evolving patterns over scale (size) and time (age).*

However, the impact of these two factors—scale and time—on network structures manifests differently. In the technology domain, time did not show any significant impact on team structure, and in some domains such as entertainment, the scale did not show a significant impact on some of the team structure measures. Generally, the variation in network characteristics due to the aging of artifacts seems to be more uniform and less category-dependent as explained before. It should also be noted that the robustness of the findings of this article is validated by testing lower time thresholds when calculating temporal

weights of all the pairs of collaborators, using 24 h instead of 48 h, we observed consistent patterns in the data. This analysis underscores the stability of our results, even when subjected to more stringent temporal constraints. For additional details on this examination, readers are directed to the last section of the Appendix document.

Our study's findings robustly validate Hypothesis 3, highlighting distinct collaborative network structures in digital platforms compared to traditional organizational settings. In this digital environment, we observed unique patterns of interconnectedness, clustering, and decentralization, contrasting with the more hierarchical structures prevalent in traditional organizations. The influence of artifact categories on network dynamics in digital contexts was particularly notable, underlining the importance of the nature of the artifact in shaping collaborative patterns. This contrast confirms the distinctiveness of digital collaboration, emphasizing the adaptability and specificity of digital platforms in facilitating collaborative processes.

*The key finding regarding Hypothesis 3: Team structure patterns in pure digital platforms drastically differ from traditional organizations.*

### C. Network Dynamics and the Category of the Artifact

In the realm of network analysis within organizational contexts, Jacobs and Watts have posited that "Network Heterogeneity Is Not Explained by Industry" [72]. This perspective suggests that the categorization of firms by industry may be too broad a metric to significantly impact the network structures within them. Building upon this viewpoint, our research pivots to a more granular level of analysis, focusing on the nature of the artifact itself. This approach is driven by the assumption that the specific category of a topic might have a more direct influence on the structural characteristics of collaborative networks.

Our investigation, centered on the digital collaborative platform of Wikipedia, as detailed in the preceding sections, reveals a clear picture of how these network characteristics unfold across different topic domains. This exploration is vital, considering that the unique nature of each artifact category could potentially foster distinct collaborative patterns.

The results of our analysis present a fascinating narrative. The scaling of teams across different content categories reveals diverse and complex patterns in team structure. Conflict and disaster categories show a consistent trend toward more cohesive and structured collaboration with distributed leadership in larger teams. In contrast, politics and technology exhibit a more dynamic evolutionary process, with initial dispersal followed by increased cohesion and subcommunity formation as teams grow. The entertainment domain, however, stands out with its more fluid and less predictable patterns, reflecting the adaptive nature of collaboration in this field. These insights offer a nuanced understanding of how team size influences collaboration, which is crucial for effectively managing and facilitating team dynamics in various contexts.

Similarly, when considering the age of artifacts, the conflicts and disasters categories again demonstrate a significant relationship with network characteristics. This could be

attributed to the time-sensitive and evolving nature of these topics, where prolonged events necessitate and foster a deepening of collaborative networks over time. On the other hand, categories such as politics and entertainment display a less consistent pattern, and the technology domain shows no meaningful pattern temporally, suggesting a multifaceted interplay of various factors influencing collaboration as these artifacts mature.

*The key finding regarding Hypothesis 1: Problem context matters in the team structure.*

## V. CONCLUSION

In an era marked by rapid technological advancement and globalization, social dynamics, facilitated by technology, adapt, and evolve within sociotechnical systems. This adaptation is particularly evident in digital collaborative environments, where technology not only supports but also transforms how knowledge is created and shared. In our study, we implemented a novel approach to network construction, tailored to capture the intricate dynamics of collaboration within digital platforms. This methodology stands out due to its innovative integration of both temporal and content-based interactions among users. By systematically analyzing edit histories across various time windows and evaluating user interactions not just in terms of alternating interactions restricted by a time threshold (ATI) but also content overlap, our approach offers a more multidimensional perspective on collaborative behaviors. This method allows for a deeper understanding of how digital collaborations evolve over time and how content-related factors influence the formation and dynamics of collaborative networks. Our study, on the backdrop of an increasingly digitalized organizational landscape, reveals critical insights into the dynamics of digital collaboration, pertinent to an era where teamwork transcends traditional communication methods.

The findings corroborate three central hypotheses, each shedding light on distinct aspects of digital team behaviors. First, validating Hypothesis 1, we found that the context of a problem significantly influences team structures on digital platforms. This underlines the importance of considering the specific nature of tasks or topics when analyzing collaborative dynamics, which is a main factor in shaping how teams form and function in digital environments. In addressing Hypothesis 2, our study unveils that team structures on digital platforms reveal evolving patterns over time and scale. The insights gained here are instrumental for understanding how digital teams grow and evolve, providing key considerations for managing and optimizing these evolving networks. Hypothesis 3 finds strong validation in our analysis, indicating that team structure patterns in purely digital platforms differ markedly from those in traditional organizational settings. This finding is crucial in an era where digital platforms are increasingly harnessed for diverse purposes, from KS to collective product and service design. Understanding these unique digital collaboration dynamics is imperative for high-level leaders and decision-makers of such platforms.

### A. Implications for Policymakers

Our findings reveal that artifact-focused peer production of knowledge fosters a more collaborative environment, particularly in larger teams. This has significant implications for decision-makers and policymakers. As teams grow in size, the potential for decentralized collaboration increases, presenting both opportunities and challenges for managing information flow and team dynamics. Policymakers should leverage the benefits of large, decentralized teams by designing structures that enhance collaborative potential and align with organizational goals. A critical consideration is the emergence of numerous subcommunities within large online teams. These subcommunities can contribute to innovation and specialization, but they also necessitate careful governance to ensure cohesive and efficient information flow.

Moreover, as our results indicate, larger teams tend to exhibit higher average degrees of interaction and increased decentralization. While this can accelerate information dissemination and decision-making, it also raises the potential for information overload and coordination challenges. Policymakers must be aware of these dynamics and implement strategies to balance speed with accuracy and cohesion. For very large teams, an integration of top-down and decentralized leadership may be a useful option to manage the extra decentralized dynamics that can emerge. Such centralized interventions are crucial for aligning team dynamics with broader organizational goals, for efficient coordination. In practice, this could involve creating roles or mechanisms for coordinating between subcommunities, setting guidelines for communication and collaboration, and utilizing technology to monitor and facilitate efficient information flow. By understanding and harnessing the dynamics of decentralized, large-scale collaboration, organizations can optimize their teams' performance and innovation capacity, thereby achieving their strategic objectives more effectively.

Last but not least, the findings highlight that the way network behavior scales is significantly dependent on the problem type, as observed in the distinct patterns. Therefore, platform governors must recognize that what is considered normal or optimal can vary based on the content and context of the problem being addressed. Tailoring governance strategies to the specific type of problem content can help in better managing and guiding collaborative efforts, ensuring that the network dynamics are conducive to achieving desired outcomes.

In conclusion, this research serves as a foundational step toward a deeper understanding of digital team interactions. As the landscape of team collaboration continues to evolve, the insights from this study will be invaluable for navigating the complexities of sociotechnical systems in the digital era. The study's insights are pivotal for organizations embracing digital platforms, guiding them toward more effective team structures and enhanced creativity. The findings not only enrich the academic discourse on digital collaboration but also offer practical applications for organizations striving to harness the full potential of digital collaborative environments in achieving their strategic goals and fostering innovation in an increasingly online world.

### VI. LIMITATIONS AND FUTURE DIRECTIONS

Although our study offers significant insights into the collaborative dynamics within Wikipedia, it is important to acknowledge its limitations and propose research paths that could further deepen our understanding of digital collaboration. The reliability of capturing collaboration evolution by the proposed network construction model is bounded in two main aspects.

- 1) Although we expect the model to be robust for a wide range of user engagement rates on Wikipedia articles, the time thresholds for counting interaction weights among users can be tuned by defining dynamic adaptive parameters. Specifically, when the speed of revising a webpage is too low (or too high), the time threshold should be adjusted to larger (or smaller) values.
- 2) The proposed methodology focuses on the main co-creation platform and does not account for back-channel communications, particularly when a subset of users coordinates separately and one of them takes responsibility for the actual edit. These are probably the most important aspects for future enhancements.

This study is focused on a single open-source digital platform suggesting the need for broader explorations. Another promising direction for future research is a cross-platform comparative analysis. This could determine if the distinctive collaborative behaviors identified on Wikipedia hold true in other digital contexts, characterized by their own interfaces, user bases, and objectives, providing a broader perspective across various digital environments.

Another crucial path for future research lies in the behavioral aspects of digital collaboration. In the digital collaboration landscape, as [6] emphasizes, trust is foundational to effective collaboration and KS. Positive team interactions have been shown to enhance trust levels. Similarly, the concept of collective psychological ownership, as introduced by Gray and colleagues, is increasingly recognized as integral to the success and creativity of teams [94], especially in digital settings where a sense of shared responsibility and investment is crucial. These factors are both supported by digital platforms' feature of traceability, where every step of the knowledge creation and collaboration process is documented. Future research should focus on how digital platforms can optimize psychological elements, and in turn, how these influence team effectiveness and innovation capacity.

### REFERENCES

- [1] M. Mosleh, P. Ludlow, and B. Heydari, "Distributed resource management in systems of systems: An architecture perspective," *Syst. Eng.*, vol. 19, no. 4, pp. 362–374, 2016.
- [2] S. Wang, W. Ding, J. Li, Y. Yuan, L. Ouyang, and F.-Y. Wang, "Decentralized autonomous organizations: Concept, model, and applications," *IEEE Trans. Comput. Social Syst.*, vol. 6, no. 5, pp. 870–878, Oct. 2019.
- [3] S. Park, T. J. Grosser, A. A. Roebuck, and J. E. Mathieu, "Understanding work teams from a network perspective: A review and future research directions," *J. Manage.*, vol. 46, no. 6, pp. 1002–1028, 2020.
- [4] L. Larson and L. A. DeChurch, "Leading teams in the digital age: Four perspectives on technology and what they mean for leading teams," *Leadership Quart.*, vol. 31, no. 1, 2020, Art. no. 101377.



- [5] N. Maddah and B. Heydari, "Building back better: Modeling decentralized recovery in sociotechnical systems using strategic network dynamics," *Rel. Eng. Syst. Saf.*, vol. 246, Jun. 2024, Art. no. 110085.
- [6] D. I. Sensuse, P. I. Lestari, and S. Al Hakim, "Exploring factors influencing knowledge sharing mechanisms and technology to support the collaboration ecosystem: A review," *DESIDOC J. Library Inf. Technol.*, vol. 41, no. 3, 2021.
- [7] S. R. Barley, B. A. Bechky, and F. J. Milliken, "The changing nature of work: Careers, identities, and work lives in the 21st century," *Academy of Management Discoveries*, vol. 3, no. 2, pp. 111–115, 2017.
- [8] B. Heydari and M. J. Pennock, "Guiding the behavior of sociotechnical systems: The role of agent-based modeling," *Syst. Eng.*, vol. 21, no. 3, pp. 210–226, 2018.
- [9] B. Heydari, Z. Szajnfarber, J. Panchal, M.-A. Cardin, K. Holtta-Otto, and G. E. Kremer, "Analysis and design of sociotechnical systems," *J. Mech. Des.*, vol. 142, no. 12, 2020, Art. no. 121101.
- [10] L. L. Gilson, M. T. Maynard, N. C. Jones Young, M. Vartiainen, and M. Hakonen, "Virtual teams research: 10 years, 10 themes, and 10 opportunities," *J. Manage.*, vol. 41, no. 5, pp. 1313–1337, 2015.
- [11] L. Lee-Kelley and N. Turner, "PMO managers' self-determined participation in a purposeful virtual community-of-practice," *Int. J. Project Manage.*, vol. 35, no. 1, pp. 64–77, 2017.
- [12] V. S. Gunasekera and S. C. Chong, "Knowledge management critical success factors and project management performance outcomes in major construction organisations in Sri Lanka: A case study," *VINE J. Inf. Knowl. Manage. Syst.*, vol. 48, no. 4, pp. 537–558, 2018.
- [13] O.-K. Choi and E. Cho, "The mechanism of trust affecting collaboration in virtual teams and the moderating roles of the culture of autonomy and task complexity," *Comput. Human Behav.*, vol. 91, pp. 305–315, Feb. 2019.
- [14] S. Gao, Y. Guo, J. Chen, and L. Li, "Factors affecting the performance of knowledge collaboration in virtual team based on capital appreciation," *Inf. Technol. Manage.*, vol. 17, pp. 119–131, Sep. 2016.
- [15] F. Scognamiglio, A. Sancino, F. Caló, C. Jacklin-Jarvis, and J. Rees, "The public sector and co-creation in turbulent times: A systematic literature review on robust governance in the COVID-19 emergency," *Public Admin.*, vol. 101, no. 1, pp. 53–70, 2023.
- [16] M. Natarajan, "Knowledge sharing through intranet," *DESIDOC J. Library Inf. Technol.*, vol. 28, no. 5, 2008, Art. no. 5.
- [17] S. B. Rahayuab, S. F. Aslah, A. M. Azahari, A. Ahmad, and N. D. Kamarudin, "Preliminary study: Knowledge sharing in collaborative e-commerce," in *Proc. 8th Int. Conf. Inf. Technol. Multimedia (ICIMU)*, Piscataway, NJ, USA: IEEE Press, 2020, pp. 78–83.
- [18] H. Suto and P. Patitad, "A representation model of collaboration in design process," in *Proc. 10th Asian Control Conf. (ASCC)*, Piscataway, NJ, USA: IEEE Press, 2015, pp. 1–5.
- [19] S. E. Page, "Making the difference: Applying a logic of diversity," *Acad. Manage. Perspect.*, vol. 21, no. 4, pp. 6–20, 2007.
- [20] T. H. Cox and S. Blake, "Managing cultural diversity: Implications for organizational competitiveness," *Acad. Manage. Perspect.*, vol. 5, no. 3, pp. 45–56, 1991.
- [21] P. B. Paulus and B. A. Nijstad, *Group Creativity: Innovation Through Collaboration*. London, U.K.: Oxford Univ. Press, 2003.
- [22] J. R. Katzenbach and D. K. Smith, *The Wisdom of Teams: Creating the High-Performance Organization*. Boston, MA, USA: Harvard Bus. Rev. Press, 2015.
- [23] M. Soltaninejad, R. Aghazadeh, S. Shaghghi, and M. Zarei, "Using machine learning techniques to forecast Mehram company's sales: A case study," *J. Bus. Manage. Stud.*, vol. 6, no. 2, pp. 42–53, 2024.
- [24] V. Ramaswamy and K. Ozcan, "What is co-creation? An interactional creation framework and its implications for value creation," *J. Bus. Res.*, vol. 84, pp. 196–205, 2018.
- [25] K. Blomqvist and J. Levy, "Collaboration capability—A focal concept in knowledge creation and collaborative innovation in networks," *Int. J. Manage. Concepts Philosophy*, vol. 2, no. 1, pp. 31–48, 2006.
- [26] K. Das and R. Mahapatra, "Collaboration and knowledge sharing via facebook groups: Analysis of postings of library and information science community," *DESIDOC J. Library Inf. Technol.*, vol. 38, no. 2, 2018, Art. no. 95.
- [27] B. R. Babu and S. Gopalakrishnan, "Knowledge sharing tools and technology: An overview," *DESIDOC J. Library Inf. Technol.*, vol. 28, no. 5, 2008, Art. no. 19.
- [28] C. B. Gibson, P. D. Dunlop, A. Majchrzak, and T. Chia, "Sustaining effectiveness in global teams: The coevolution of knowledge management activities and technology affordances," *Org. Sci.*, vol. 33, no. 3, pp. 1018–1048, 2022.
- [29] Y. Charband and N. Jafari Navimipour, "Online knowledge sharing mechanisms: A systematic review of the state of the art literature and recommendations for future research," *Inf. Syst. Frontiers*, vol. 18, pp. 1131–1151, Feb. 2016.
- [30] C. B. Gibson and J. L. Gibbs, "Unpacking the concept of virtuality: The effects of geographic dispersion, electronic dependence, dynamic structure, and national diversity on team innovation," *Admin. Sci. Quart.*, vol. 51, no. 3, pp. 451–495, 2006.
- [31] S. L. Marlow, C. N. Lacerenza, and E. Salas, "Communication in virtual teams: A conceptual framework and research agenda," *Human Res. Manage. Rev.*, vol. 27, no. 4, pp. 575–589, 2017.
- [32] H. Zhu, R. Kraut, and A. Kittur, "Effectiveness of shared leadership in online communities," in *Proc. ACM Conf. Comput. Supported Cooperative Work*, 2012, pp. 407–416.
- [33] H. Zhu, R. E. Kraut, and A. Kittur, "Effectiveness of shared leadership in wikipedia," *Human Factors*, vol. 55, no. 6, pp. 1021–1043, 2013.
- [34] A. Kittur and R. E. Kraut, "Harnessing the wisdom of crowds in wikipedia: quality through coordination," in *Proc. ACM Conf. Comput. Supported Cooperative Work*, 2008, pp. 37–46.
- [35] A. Kittur and R. E. Kraut, "Beyond wikipedia: coordination and conflict in online production groups," in *Proc. ACM Conf. Comput. Supported Cooperative Work*, 2010, pp. 215–224.
- [36] A. Peña-Calvin, J. Saldivar, J. Arroyo, and S. Hassan, "A categorization of decentralized autonomous organizations: the case of the Aragon platform," *IEEE Trans. Comput. Social Syst.*, early access, 2023.
- [37] H. W. Chesbrough and M. M. Appleyard, "Open innovation and strategy," *California Manage. Rev.*, vol. 50, no. 1, pp. 57–76, 2007.
- [38] J. McGrath, "Leadership behavior: some requirements for leadership training united states civil service commission," *Office of Career Development*, US Civil Service Commission, 1962.
- [39] S. J. Zaccaro, A. L. Rittman, and M. A. Marks, "Team leadership," *Leadership Quart.*, vol. 12, no. 4, pp. 451–483, 2001.
- [40] F. P. Morgeson, D. S. DeRue, and E. P. Karam, "Leadership in teams: A functional approach to understanding leadership structures and processes," *J. Manage.*, vol. 36, no. 1, pp. 5–39, 2010.
- [41] L. D'Innocenzo, J. E. Mathieu, and M. R. Kukenberger, "A meta-analysis of different forms of shared leadership—team performance relations," *J. Manage.*, vol. 42, no. 7, pp. 1964–1991, 2016.
- [42] V. C. Nicolaidis et al., "The shared leadership of teams: A meta-analysis of proximal, distal, and moderating relationships," *Leadership Quart.*, vol. 25, no. 5, pp. 923–942, 2014.
- [43] D. Wang, D. A. Waldman, and Z. Zhang, "A meta-analysis of shared leadership and team effectiveness," *J. Appl. Psychol.*, vol. 99, no. 2, 2014, Art. no. 181.
- [44] B. Heydari, M. Mosleh, and K. Dalili, "Efficient network structures with separable heterogeneous connection costs," *Econ. Lett.*, vol. 134, pp. 82–85, Sep. 2015.
- [45] M. Mosleh, K. Dalili, and B. Heydari, "Distributed or monolithic? A computational architecture decision framework," *IEEE Syst. J.*, vol. 12, no. 1, pp. 125–136, Jan. 2016.
- [46] H. Zhu, R. E. Kraut, Y.-C. Wang, and A. Kittur, "Identifying shared leadership in wikipedia," in *Proc. SIGCHI Conf. Human Factors Comput. Syst.*, 2011, pp. 3431–3434.
- [47] J. Wales, "Wikimedia," *PediaPress*, 2001.
- [48] S. Jhaver, S. Frey, and A. X. Zhang, "Decentralizing platform power: A design space of multi-level governance in online social platforms," *Social Media+ Soc.*, vol. 9, no. 4, 2023, Art. no. 20563051231207857.
- [49] N. T. Korfiatis, M. Poulos, and G. Bokos, "Evaluating authoritative sources using social networks: An insight from wikipedia," *Online Inf. Rev.*, vol. 30, no. 3, pp. 252–262, 2006.
- [50] U. Brandes, P. Kenis, J. Lerner, and D. Van Raaij, "Network analysis of collaboration structure in wikipedia," in *Proc. 18th Int. Conf. World Wide Web*, 2009, pp. 731–740.
- [51] D. Laniado, R. Tasso, Y. Volkovich, and A. Kaltenbrunner, "When the wikipediaians talk: Network and tree structure of wikipedia discussion pages," in *Proc. Int. AAAI Conf. Web Social Media*, vol. 5, no. 1, 2011, pp. 177–184.
- [52] S. Maniu, B. Cautis, and T. Abdessalem, "Building a signed network from interactions in wikipedia," in *Databases and Social Networks*, New York, NY, USA: Association for Computing Machinery, 2011, pp. 19–24.

- [53] T. Yasserli, R. Sumi, A. Rung, A. Kornai, and J. Kertész, "Dynamics of conflicts in wikipedia," *PLoS One*, vol. 7, no. 6, 2012, Art. no. e38869.
- [54] D. M. Lydon-Staley, D. Zhou, A. S. Blevins, P. Zurn, and D. S. Bassett, "Hunters, busybodies and the knowledge network building associated with deprivation curiosity," *Nature Human Behav.*, vol. 5, no. 3, pp. 327–336, 2021.
- [55] K. Zhu, D. Walker, and L. Muchnik, "Content growth and attention contagion in information networks: Addressing information poverty on wikipedia," *Inf. Syst. Res.*, vol. 31, no. 2, pp. 491–509, 2020.
- [56] E. Bassani and M. Viviani, "Automatically assessing the quality of wikipedia contents," in *Proc. 34th ACM/SIGAPP Symp. Appl. Comput.*, 2019, pp. 804–807.
- [57] P. Wang and X. Li, "Assessing the quality of information on wikipedia: A deep-learning approach," *J. Assoc. Inf. Sci. Technol.*, vol. 71, no. 1, pp. 16–28, 2020.
- [58] Y. Ren, H. Zhang, and R. E. Kraut, "How did they build the free encyclopedia? A literature review of collaboration and coordination among wikipedia editors," *ACM Trans. Comput.-Human Interact.*, vol. 31, no. 1, pp. 1–48, 2023.
- [59] A.-M. Palos, "The information flow in virtual teams," *J. Adv. Res. Manage. (JARM)*, vol. 3, no. 5, pp. 38–45, 2012.
- [60] N. Maddah and E. Roghanian, "Data-driven performance management of business units using process mining and dea: Case study of an Iranian chain store," *Int. J. Product. Perform. Manage.*, vol. 72, no. 2, pp. 550–575, 2023.
- [61] D. J. Watts and S. H. Strogatz, "Collective dynamics of small-world networks," *Nature*, vol. 393, no. 6684, pp. 440–442, 1998.
- [62] R. S. Burt, "Structural holes and good ideas," *Amer. J. Sociol.*, vol. 110, no. 2, pp. 349–399, 2004.
- [63] B. Uzzi and J. Spiro, "Collaboration and creativity: The small world problem," *Amer. J. Sociol.*, vol. 111, no. 2, pp. 447–504, 2005.
- [64] J. Singh, M. T. Hansen, and J. M. Podolny, "The world is not small for everyone: Inequity in searching for knowledge in organizations," *Manage. Sci.*, vol. 56, no. 9, pp. 1415–1438, 2010.
- [65] D. A. Gianetto and B. Heydari, "Network modularity is essential for evolution of cooperation under uncertainty," *Scientific Rep.*, vol. 5, no. 1, 2015, Art. no. 9340.
- [66] K. Yang, I. Fujisaki, and K. Ueda, "Cooperation patterns of members in networks during co-creation," *Scientific Rep.*, vol. 11, no. 1, 2021, Art. no. 11588.
- [67] D. Lazer and A. Friedman, "The network structure of exploration and exploitation," *Admin. Sci. Quart.*, vol. 52, no. 4, pp. 667–694, 2007.
- [68] W. A. Mason, A. Jones, and R. L. Goldstone, "Propagation of innovations in networked groups," *J. Exp. Psychol., General*, vol. 137, no. 3, 2008, Art. no. 422.
- [69] H. Zhu, Y. Sheng, X. Zhou, and Y. Zhu, "Group role assignment with cooperation and conflict factors," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 48, no. 6, pp. 851–863, Jun. 2016.
- [70] H. Zhu, "Avoiding conflicts by group role assignment," *IEEE Trans. Syst., Man, Cybern.: Syst.*, vol. 46, no. 4, pp. 535–547, Apr. 2015.
- [71] Q. Jiang, H. Zhu, Y. Qiao, D. Liu, and B. Huang, "Extending group role assignment with cooperation and conflict factors via kd45 logic," *IEEE Trans. Comput. Social Syst.*, vol. 10, no. 1, pp. 178–191, Jan. 2022.
- [72] A. Z. Jacobs and D. J. Watts, "A large-scale comparative study of informal social networks in firms," *Manage. Sci.*, vol. 67, no. 9, pp. 5489–5509, 2021.
- [73] G. Kossinets and D. J. Watts, "Origins of homophily in an evolving social network," *Amer. J. Sociol.*, vol. 115, no. 2, pp. 405–450, 2009.
- [74] L. Yang et al., "The effects of remote work on collaboration among information workers," *Nature Human Behav.*, vol. 6, no. 1, pp. 43–54, 2022.
- [75] M. Chen, K. Kuzmin, and B. K. Szymanski, "Community detection via maximization of modularity and its variants," *IEEE Trans. Comput. Social Syst.*, vol. 1, no. 1, pp. 46–65, Jan. 2014.
- [76] J. L. Gibbs, A. Sivunen, and M. Boyraz, "Investigating the impacts of team type and design on virtual team processes," *Human Resour. Manage. Rev.*, vol. 27, no. 4, pp. 590–603, 2017.
- [77] R. Reagans and E. W. Zuckerman, "Networks, diversity, and productivity: The social capital of corporate R&D teams," *Org. Sci.*, vol. 12, no. 4, pp. 502–517, 2001.
- [78] R. Reagans, E. Zuckerman, and B. McEvily, "How to make the team: Social networks vs. demography as criteria for designing effective teams," *Admin. Sci. Quart.*, vol. 49, no. 1, pp. 101–133, 2004.
- [79] D. Obstfeld, "Social networks, the tertius iungens orientation, and involvement in innovation," *Admin. Sci. Quart.*, vol. 50, no. 1, pp. 100–130, 2005.
- [80] B. Kogut and G. Walker, "The small world of Germany and the durability of national networks," *Amer. Sociol. Rev.*, vol. 66, no. 3, pp. 317–335, 2001.
- [81] D. Centola and M. Macy, "Complex contagions and the weakness of long ties," *Amer. J. Sociol.*, vol. 113, no. 3, pp. 702–734, 2007.
- [82] W. Mason and D. J. Watts, "Collaborative learning in networks," *Proc. Nat. Acad. Sci.*, vol. 109, no. 3, pp. 764–769, 2012.
- [83] E. Muller and R. Peres, "The effect of social networks structure on innovation performance: A review and directions for research," *Int. J. Res. Marketing*, vol. 36, no. 1, pp. 3–19, 2019.
- [84] S. P. Borgatti and P. C. Foster, "The network paradigm in organizational research: A review and typology," *J. Manage.*, vol. 29, no. 6, pp. 991–1013, 2003.
- [85] P. R. Monge and N. S. Contractor, *Theories of Communication Networks*. London, U.K.: Oxford Univ. Press, USA, 2003.
- [86] N. Katz, D. Lazer, H. Arrow, and N. Contractor, "Network theory and small groups," *Small Group Res.*, vol. 35, no. 3, pp. 307–332, 2004.
- [87] S. Sasidharan, R. Santhanam, D. J. Brass, and V. Sambamurthy, "The effects of social network structure on enterprise systems success: A longitudinal multilevel analysis," *Inf. Syst. Res.*, vol. 23, no. 3-part-1, pp. 658–678, 2012.
- [88] J. Becker, D. Brackbill, and D. Centola, "Network dynamics of social influence in the wisdom of crowds," *Proc. Nat. Acad. Sci.*, vol. 114, no. 26, pp. E5070–E5076, 2017.
- [89] J. Leskovec, J. Kleinberg, and C. Faloutsos, "Graph evolution: Densification and shrinking diameters," *ACM Trans. Knowl. Discovery Data (TKDD)*, vol. 1, no. 1, p. 2-es, 2007.
- [90] H. A. Simon, "The architecture of complexity," *Proc. Amer. Philosoph. Soc.*, vol. 106, no. 6, pp. 467–482, 1962.
- [91] S. Padhee, N. Lore, and B. Heydari, "Design teams and industry life cycles: The interplay of innovation and complexity," *Syst. Eng.*, vol. 26, no. 5, pp. 604–617, 2023.
- [92] P. Holme and B. J. Kim, "Growing scale-free networks with tunable clustering," *Phys. Rev. E*, vol. 65, no. 2, 2002, Art. no. 026107.
- [93] M. E. Newman, "The structure and function of complex networks," *SIAM Rev.*, vol. 45, no. 2, pp. 167–256, 2003.
- [94] S. M. Gray, A. P. Knight, and M. Baer, "On the emergence of collective psychological ownership in new creative teams," *Org. Sci.*, vol. 31, no. 1, pp. 141–164, 2020.



**Negin Maddah** received the B.S. degree in industrial engineering from Alzahra University, Tehran, Iran, in 2017, and the M.S. degree in industrial engineering, systems optimizations from K. N. Toosi University of Technology, Tehran, in 2020. She is currently working toward the Ph.D. degree in industrial engineering with the Multi-Agent Intelligent Complex Systems (MAGICS) Lab, Northeastern University, Boston, MA, USA.

Her research interests include computational social sciences, sociotechnical systems, platform-based systems, process mining, and data mining.



**Babak Heydari** received the B.S. degree in electrical engineering from Sharif University of Technology, Tehran, Iran, in 2002, and the M.S. and Ph.D. degrees in electrical engineering and computer science from the University of California, Berkeley, CA, USA, in 2005 and 2008, respectively.

He is an Associate Professor with the College of Engineering, Northeastern University, Boston, MA, USA. He holds affiliated positions with the Network Science Institute, the Institute for Experiential AI, and the School of Public Policy, Northeastern University. He also serves as the Director of the Multi-Agents Intelligent Complex Systems (MAGICS) Laboratory and the Engineering Management Program, Northeastern University. His research interests include network science, the design and governance of human-AI systems, computational social science, and multisided platforms.

Dr. Heydar was a recipient of the NSF CAREER Award.